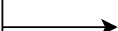
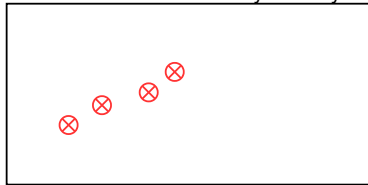


Observed Spatial Dynamics

$(x_{\text{center}}, y_{\text{center}}, v_x, v_y, a_x, a_y)$



GRU-1
Encoder

h_t^{e1}



Concatenation



h_t^c

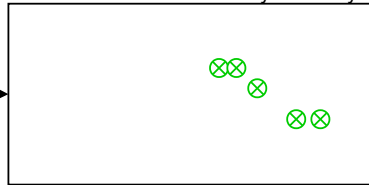


GRU-3
Decoder



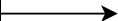
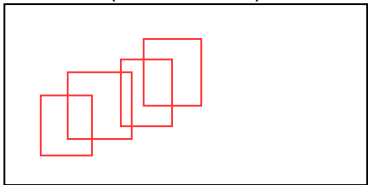
Predicted Spatial Dynamics

$(x'_{\text{center}}, y'_{\text{center}}, v'_x, v'_y, a'_x, a'_y)$



Observed Dimensional Attributes

$(w, h, \Delta w, \Delta h)$



GRU-2
Encoder

h_t^{e2}



GRU-4
Decoder



Predicted Dimensional Attributes

$(w', h', \Delta w', \Delta h')$

